ROUTING AND RECORD SHEET

SUBJECT: (Optional)

Proposed Publication of Research Report

FILE *Printing & Reproduction*

FROM: PSS/OMS Rm. 412 1000 Glebe	EXTENSION	NO. DD/S 71-2129
		DATE 11 May 1971

STAT

TO: (Officer designation, room number, and building)	DATE RECEIVED	DATE FORWARDED	OFFICER'S INITIALS	COMMENTS (Number each comment to show from whom to whom. Draw a line across column after each comment.)
1. C/PSS/OMS 1000 Glebe		26 May 71		I have reviewed the attached research report and I find it consistent with the policy guidance governing unofficial open publication by CIA employees as set forth in _____ Publication of this report in the journal cited would constitute a significant contribution to the literature in this field. It is also a good example of the type of test validation research called for by the guidelines issued by the Equal Opportunity Commission and affirmed by a recent Supreme Court ruling.
2.				
3.				
4.				
5.				
6.				
7.				
8. D/MS 1D-4061 Hqs.	26 MAY 1971	26 MAY 1971		#8 - Recommend approval.
9. Deputy Director for Support - 7D-26 Hqs.				C/PSS/OMS
10. Director of Security 4E-60 Hqs.	2 June 71		Rob	to 10 and 11 and back to 12 for approval
11. Asst to the Director 1F-04 Hqs.	4 June 71			
12. Deputy Director for Support Rm 7D18, Hqs	9 JUN 1971			App_____
13. Director of Medical Services Rm 1D4060, Hqs				
14.				
15.				

STAT (multiple)

Robert S. Wattles
Assistant Deputy Director
for Support   9 JUN 1971

DD/S Distribution:
Orig - DMS w/Orig of DD/S 71-2129
1 - DD/S Chrono
1 - DD/S Subject, w/cy of 71-2129

11 May 1971

MEMORANDUM FOR: Director of Medical Services

THROUGH   : Chief, Psychological Services Staff

SUBJECT   : Proposed Publication of Research
                  Report

　　　1. Attached is a copy of a research report I would like
to submit for publication to Personnel Psychology, a journal
of applied research in psychology. Subject to your approval,
it will be formally submitted to the Office of Security and
to the Assistant to the DCI for their approval.

　　　2. The attached report details the first validation
study of a widely-used test designed to select computer
programmers. Specifically, the study demonstrated that
the AABP test (Aptitude Assessment Battery: Programming)
is significantly predictive of both training and job per-
formance of programmers. It is believed that these find-
ings will be of interest to all who use this test, as well
as to those who have considered its use but are reluctant
to do so without some evidence of its efficacy as a measure
of programming aptitude.

　　　3. Thank you very much for your consideration of this
report.

25X1

Acting Chief, Research Branch
Psychological Services Staff
Office of Medical Services

SECRET

28 May 1971

25X1

NOTE FOR: Mr. Coffey via Mr. Wattles via

    Recommend your approval. Validation
studies were undertaken at the request, and
with the cooperation of OCS.

    There is a credit in footnote for
the Chief of the OCS Training Program.

25X1

25X1

1 June 71

Mr. W. -

Even though para 1 of the
attached does not indicate it,
this appears to be the "formal"
request for approval.

If you agree, I will add a
"DD/S approval" line. Also,
although Dr. Tietjen's
"recommendation" is on the
RS, I would think the approval
might be more appropriate
on Dr. [   ] memo?????

STAT

p.

I have attached a cy of
[   ] for your ref.

STAT

11 May 1971

MEMORANDUM FOR: Director of Medical Services

THROUGH   : Chief, Psychological Services Staff

SUBJECT   : Proposed Publication of Research
               Report

1. Attached is a copy of a research report I would like to submit for publication to Personnel Psychology, a journal of applied research in psychology. Subject to your approval, it will be formally submitted to the Office of Security and to the Assistant to the DCI for their approval.

2. The attached report details the first validation study of a widely-used test designed to select computer programmers. Specifically, the study demonstrated that the AABP test (Aptitude Assessment Battery: Programming) is significantly predictive of both training and job performance of programmers. It is believed that these findings will be of interest to all who use this test, as well as to those who have considered its use but are reluctant to do so without some evidence of its efficacy as a measure of programming aptitude.

3. Thank you very much for your consideration of this report.

                        Acting Chief, Research Branch
                        Psychological Services Staff
                        Office of Medical Services

25X1

SECRET
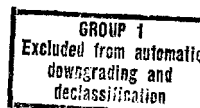
PREDICTION OF COMPUTER PROGRAMMER TRAINING
AND JOB PERFORMANCE USING THE AABP TEST[1]

STAT

Central Intelligence Agency

The rapid growth of the field of computer programming in the past ten years has generated a need to identify individuals with the talents to profit from training and become productive programmers. This need for early identification of potential programmers has been approached in two main ways by personnel psychologists: 1) creation of special tests designed to measure the specific aptitudes believed essential to computer programming, and 2) application of pre-existing tests. Unfortunately, the number of published validation studies which have resulted from the use of both types of tests seems disproportionately small in relation to the massive use of these tests as programmer selection devices for government, industry, and specialized training schools.

Of the several aptitude tests specifically designed to select potential programmers, various versions of the IBM Programmers Aptitude Test have been most frequently subjected to controlled validity studies. In nearly all instances, these studies reported relationships between test scores and <u>training</u> performance. Significant correlations

between the Programmers Aptitude Test and programmer train-
ing have been reported by McNamara and Hughes (1961), Oliver
and Willis (1963), Katz (1964), Hollenbeck and McNamara (1965)
and Bauer, Mehrens, and Vinsonhaler (1968). Of these studies,
only McNamara and Hughes (1961) conducted research relating
test scores to actual job performance; they reported correla-
tions of .44 and .36 between Programmers Aptitude Test Scores
and supervisors' ratings of job performance.

Other specially-designed aptitude tests have shown at
least some promise for predicting programmer training per-
formance, including the Logical Analysis Device (McNamara
and Hughes, 1961), the Computer Usage Company Programmer
Aptitude Test (Hollenbeck and McNamara, 1965) and the
Computer Programmer Aptitude Battery (Perry, 1967a). More
general purpose tests which have been successfully used to
predict programmer training performance include the Army's
General Technical Aptitude Area (Katz, 1962), the Wonderlic
Personnel Test (Biamonte, 1965), the mathematics test from
the Navy Officer Classification Battery (Myer, 1965), the
College Qualification Test (Bauer, Mehrens and Vinsonhaler,
1968) and the Strong Vocational Interest Blank (Bauer, Mehrens
and Vinsonhaler, 1968). As with the IBM Programmers Aptitude
Test, reports of relationships between test and job perfor-
mance for these other tests are considerably scarcer.

STAT
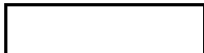
Seiler (1965) reported correlations ranging from .35 to .62 between supervisor's ratings of workers in electronic data processing jobs and combinations of three or more tests from the General Aptitude Test Battery, and Perry (1967b) found that an occupational key constructed using SVIB procedures was related to job satisfaction but not to relative salary of computer programmers.

One widely-used test of programming aptitude -- the Aptitude Assessment Battery: Programming (AABP) -- is absent in the validity literature. Despite its lack of published validity research, its author claims that is has been administered to over 3600 persons in 23 different companies, institutions and government agencies (Wolfe, 1970). According to the brochure published by Programming Specialists, Incorporated, which describes the AABP, it is an untimed test which attempts to evaluate the applicant's aptitude for protracted concentration on a long sequence of steps. Both the process of problem-solving as well as the answers to the items form the basis for evaluating individual performance. The scoring and interpretation of the AABP is typically performed by the distributor of the test.

The purpose of the present paper is to report the relationships between programmer aptitude as measured by the AABP and both the training and job performance of computer

STAT [   ]

4

programmers in a ~~large~~ governmental agency. It is antici-
pated that these results will be of interest to those who
are using this test without benefit of any validity data,
as well as to those who may be considering its use but are
reluctant to do so without some evidence of its efficacy
as a measure of programming aptitude.

Method

Measurement of Training Performance

Ninety-three students, 71 men and 22 women, from seven
runnings of an inhouse computer programming course were se-
lected for study. The average age of these students was
28 years, with a range in age from 20 to 52. 4% had earned
advanced degrees, 46% possessed undergraduate degrees, 24%
had some college experience and 12% had attended one or
more business or technical schools; the remainder were high
school graduates. The course, which was 15 weeks in length,
was designed to produce professional programmers capable of
writing programs in both PL/1 and ALC programming languages.
Evaluation of students in these classes was based upon
weekly tests (which accounted for about 70% of the students'
grades) and evaluation of the programs written by the stud-
ents (which accounted for the remaining 30%). Based upon
their total grades, students were given a rank representing
their final standing among the graduates in their class.
The

The AABP was administered to the 93 students prior to their admission to the programming course. Not all students in the seven training classes took the AABP; 35 persons who completed the course had not taken this test. While the test results entered into the selection decision for many of the students in this course, these results were not known by the instructors who taught and evaluated the students.

Measurement of Programming Performance

A total of 57 programmers, 36 men and 21 women, were included in this portion of the study; none of these people were present in the training sample described above. The average age of this group was 25 years, with a range in age from 20 to 50. The number of months of programming experience of individuals in this group ranged from 1-78 with an average of slightly over 15 months. A variety of programming languages was used by these programmers in their daily work, with Fortran being most commonly used followed by PL/1, Cobol and ALC. Each subject was rated by his immediate supervisor on job performance as a programmer, future programming potential, potential for systems analysis and potential for management. All ratings were made on a 7-point, modified "letter-grade" scale, with ratings ranging from A through F. Two additional categories ("Between A and B" and "Between B and C") were added in an attempt to expand the variance of the

ratings. In nearly all cases, the programmers were rated before they were tested; thus this portion of the study was a test of the concurrent validity of the AABP.

Results

Training Performance

For each of the seven programming classes, the class rank standings assigned to the members of each class were normalized according to the normalized-rank method (Guilford, 1954). This score transformation was utilized since rankings form a perfectly rectangular distribution; a normalized distribution of rankings is more likely to approximate the real distribution of performances among members of a class. Table 1 presents the correlations between AABP scores and normalized class ranks for each of the seven programming classes. Correlations ranged from .33 to .66 with a median of .44. When all classes were combined, the correlation between AABP numerical scores and normalized class ranks was .40.

Insert Table 1 about here

When the AABP test is scored by its distributor, in addition to the numerical score a subjective adjectival rating is provided estimating the individual's overall potential for computer programming. This adjectival rating is highly related

to the numerical score on the AABP, although the rating takes
into account additional subjective factors including the ex-
aminee's "method, conception and approach which ... may indi-
cate sufficient superior aptitude to compensate for his lack
of accuracy in the less challenging problems" (from the bro-
chure of Programming Specialists, Incorporated, describing the
AABP). That the adjectival rating is highly related but not
completely equivalent to the numerical score is evidenced by
a correlation of .94 between adjectival rating and numerical
score for the overall training sample of 93 persons. Table 1
presents the correlations between these adjectival ratings and
normalized class ranks for each of the seven programming class-
es, which range from .22 to .66 with a median of .44. When
all classes were combined, the correlation between AABP ad-
jectival ratings and normalized class ranks was .39, essent-
ially the same as the .40 obtained with AABP numerical scores.

Table 2 presents the relationship between AABP adjectival
ratings and performance in programming training in expectancy
table form. Individuals in all seven classes were divided
into two groups--those who finished in the upper half of their
classes and those who finished in the lower half. At one ex-
treme, 7 out of 9 of those with the highest ratings on the
AABP finished in the top half of their classes, while at the
other extreme, only 20% of those with the lowest ratings did so.
so.

STAT

8

Inspection of Table 2 suggests that the AABP adjectival ratings were related to training outcomes in a manner consistent with the commonly-accepted meanings of these adjectives.

Insert Table 2 about here

Programming Performance

Table 3 presents the correlations between AABP results and supervisors' ratings of current performance and future potential for 57 programmers. Moderate-sized correlations, ranging from .40 to .46, were obtained between the AABP adjectival ratings and ratings of current job performance, programming potential, and systems analysis potential, with a smaller correlation (.30) emerging between the AABP and potential for management. Slightly smaller correlations were obtained between AABP numerical scores and ratings of job performance and potential, although in no instance was one of these correlations significantly smaller than the corresponding correlation between AABP adjectival ratings and ratings of job performance and potential.

Insert Table 3 about here

Discussion

The results of this study suggest that the AABP is a

reasonably effective device for selection of computer pro-
grammers. Using independent samples, AABP test scores cor-
related significantly with both programming training and job
performance as a programmer. No real evidence was obtained
supporting the tests distributor's implied hypothesis that
AABP adjectival ratings, which take into account subjective
factors not considered in the numerical scores, are superior
predictors of training or job performance of programmers.
Since the two types of scores are so highly correlated, and
since they correlate about equally well with training and job
performance, they would appear to be interchangeable.

Somewhat surprisingly, the AABP test was as highly re-
lated to job as to training performance. This finding is
in contrast to the research reported for the IBM Programmers
Aptitude Test which has produced correlations ranging from
the .30's to the .50's and .60's with training performance
(McNamara and Hughes, 1961; Katz, 1964; Hollenbeck and Mc-
Namara, 1965; Bauer, Mehrens and Vinsonhaler, 1968) but sig-
nificantly lower relationships with actual job performance
as a programmer (McNamara and Hughes, 1961). Since the
correlations obtained in this study between the AABP and
ratings of job performance and potential are comparable in
size to those reported by McNamara and Hughes (1961), it may
be tentatively concluded that AABP scores are as related to

performance as a programmer as are scores from the Program-

mers Aptitude Test. It should be mentioned again, however,

that the present study was a test of the <u>concurrent</u> validity

of the AABP with regard to job performance and potential; its

<u>predictive</u> validity remains to be established. Meanwhile,

the present study provided reassurance that the AABP -- a

widely used test despite its previous lack of any published

validity studies -- appears to be a reasonably valid measure

of computer programming aptitude.

References

Biamonte, A. J. A study of the effect of attitudes on the learning of computer programming. In M. H. Gotterer (Ed.), Proceedings of the third annual computer research conference. Silver Spring, Md.: Computer Personnel Research Group, 1965. Pp.

Guilford, J. P. Psychometric methods. (2nd ed.) New York: McGraw-Hill, 1954.

Bauer, R., Mehrens, W. A., & Vinsonhaler, J. F. Predicting performance in a computer programming course. Educational and Psychological Measurement, 1968, 28, 1159-1166.

Hollenbeck, G. P., & McNamara, W. J. CUCPAT and programming aptitude. Personnel Psychology, 1965, 18, 101-106.

Katz, A. Prediction of success in automatic data processing course. U. S. Army Personnel Research Office Technical Note, 1962, No. 126.

McNamara, W. J., & Hughes, J. L. A review of research on the selection of computer programmers. Personnel Psychology, 1961, 14, 39-51.

Meyer, J. K. An experimental comparison of instructional techniques for use in teaching computer program flow chart design. USN PRA Technical Bulletin, 1965, No. 65-10.

Oliver, T. C., & Willis, W. K. A study of the validity of the programmer aptitude test. Educational and Psycholog-

ical Measurement, 1963, 23, 823-825.

Perry, D. K. Evaluation of tests for improvement of programmer trainee selection. SDC Technical Memorandum 3570, System Development Corporation, 1967 (a).

Perry, D. K. Vocational interests and success of computer programmers. Personnel Psychology, 1967, 20, 517-524 (b).

Seiler, J. Abilities for ADP occupations. In M. H. Gotterer (Ed.), Proceedings of the third annual computer research conference. Silver Spring, Md.: Computer Personnel Research Group, 1965.
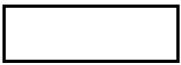
Wolfe, J. M. A new look at programming aptitudes. Business Automation, 1970, , 36-45.
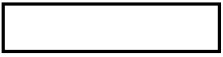
STAT

13

Footnotes

[1]The views expressed in this article are those of the author and do not necessarily reflect an official position of the Central Intelligence Agency. The author gratefully acknowledges the assistance of Mr. [   ] who provided training performance information used in the study.

[2]Requests for reprints should be sent to [   ]

STAT

STAT

STAT

[   ] Central Intelligence Agency, Washington, D.C. 20505.

STAT

Table 1

Correlations Between AABP Scores and Normalized

Class Ranks in Seven Programming Classes[-]

Class	N[a]	r with AABP Numerical Score	r with AABP Adjectival Rating
1	12	.35	.53
2	16	.45	.40
3	20	.66**	.66**
4	9	.33	.22
5	21	.44*	.41
6	5	.44	.56
7	9	.34	.44
All Classes	93	.40**	.39

Note.--For the total sample, mean AABP numerical score

was 78.2 with a standard deviation of 14.5.

[a]For most classes, the actual number of students was

larger than these N's, which represent the number of students

in each class who had taken the AABP.

**p<.01

* p<.05

STAT

15

Table 2

Relationship Between AABP Adjectival Ratings and

Instructors' Rankings in Programming Courses

Adjectival Rating	N	Percent in Upper Half of Class	Percent in Lower Half of Class
Outstanding, Excellent, Very High	9	78	22
High, Above Average	25	68	32
Average, Satisfactory S	39	49	51
Limited use, Below Average, Low	20	20	80
Totals	93	51	49

Table 3

Correlations Between AABP Scores and Supervisors'

Ratings of Performance and Potential

for 57 Programmers

Rating	N^a	r with AABP Adjectival Rating	r with AABP Numerical Score
Current Job Performance	57	.40**	.34**
Programming Potential	56	.41**	.33**
System Analysis Potential	56	.46**	.38**
Management Potential	50	.30*	.27

Note.--For the total sample, mean AABP score was 81.4 with a standard deviation of 15.1.

^{a}Some supervisors declined to rate individuals on all dimensions.

**$p < .01$

* $p < .05$

STAT

STAT

Bonnie:

The Regulation says requests to publish should come to D/S through Central Cover, Security, and Asst to the Director.

Mr. [ ] recalls that D/S has issued separate instructions requesting that they come here first. [ ] will check his records to see if he can substantiate this. Will you see if you can find out here what the precedents have been, and if there is an instruction of that sort?

W.

Will do the same as in the past. To D/OS & A-DCI & back here for approval.